SUBSTITUTE SPECIFICATION

[001] This application claims the benefit of U.S. Provisional Patent Application No. 60/311,824, filed August 14, 2001, and U.S. Provisional Patent Application No. 60/313,523, filed August 21, 2001, the disclosures of each of which are incorporated herein by reference in their entirety.

BACKGROUND OF THE INVENTION

[002] The present invention pertains to polynucleotides derived from *Mycobacterium tuberculosis* (*M. tuberculosis*) genes that aid in imparting resistance to antibiotics and chemically related compounds. This invention also relates to the use of the polynucleotides as oligonucleotide primers or probes for detecting *M. tuberculosis* strains that are resistant to antibiotics and related compounds in a biological sample. The present invention is also directed to diagnostic kits for detecting specific strains of *M. tuberculosis* expected to be contained in a biological sample.

[003] Tuberculosis remains the world's leading infectious cause of adult deaths due to a single pathogen. Outbreaks of Multidrug-Resistant (MDR) tuberculosis defined as resistant to rifampicin and isoniazid are numerous, with low rates of treatment response and very high mortality. Some of these outbreaks involve patients with HIV infection (1, 2). In some reports, strains with a particular genotype have been identified, such as the "W" strain, which caused a major outbreak in New York (3).

[004] In 1995, it was reported that the largest proportion of the *M. tuberculosis* strains from the Beijing area had highly similar IS6110 restriction fragment length polymorphism (RFLP) patterns and identical spoligo patterns (4). These strains were therefore designated "Beijing" genotype strains. The "W" variant strains in New York

appeared to represent the Beijing genotype of *M. tuberculosis*. Figure 1 depicts an example of the characteristic IS6110 RFLP and highly specific spoligo patterns of the Beijing genotype. The unexpected level of genetic conservation among *M. tuberculosis* isolates of the Beijing genotype led to the hypothesis that these strains may have selective advantages over other *M. tuberculosis* strains. Moreover, Beijing genotype strains were significantly associated with drug resistance in Cuba, Estonia, and Vietnam (vanSoolingen unpublished).

[005] Several studies have suggested that strains of the Beijing genotype are emerging. In Vietnam, the proportion of Beijing strains was 71% in patients aged under 25, and 41% for those over 55 (vanSoolingen unpublished). Beijing strains have been implicated in several tuberculosis epidemics in the U.S.A (3) and recently in Gran Canaria (5). A recent study showed that 82% of MDR strains isolated in a prison in Azerbaijan, Eastern Europe, are of the Beijing genotype (6).

[006] Knowledge of a possible molecular basis for the acquisition of drug resistance in *M. tuberculosis* would be extremely useful as a prerequisite for appropriate treatment. Resistance to antibiotics in *M. tuberculosis* is due to genomic mutations in specific genes of the bacterium. In contrast to other Gram-positive or Gram-negative pathogens with MDR phenotypes, plasmid or transposon-mediated mechanisms of resistance have not been reported in *M. tuberculosis* (3,4,5).

[007] Recent studies have provided evidence for a role of mutator phenotypes in the emergence of MDR clinical *Pseudomonas* isolates (6). Such phenotypes may enable the bacteria, not only to easily acquire resistance to several antibiotics, but also to adapt to new niches and to escape immune surveillance by modulating bacterial

resistance to host defenses (7). This prompted an investigation into whether a similar situation might exist in *M. tuberculosis*. Since MDR strains are dangerous to manipulate, it was decided, as a first step, to simply look for the presence of mutations in genes expected to play a role in mutation frequency.

[008] Mutator phenotypes commonly result from defects in DNA repair enzymes (8). An *in silico* analysis suggested that most mismatch repair systems, like *mutS, mutL* or *mutH,* are missing in the *M. tuberculosis* genome. However, the frequency of spontaneous mutations in *M. tuberculosis* (*in vitro* cultures) is similar to that found in other bacteria carrying mismatch repair systems. This characteristic could have significant implications with respect to genome stability and strain variability (9). This suggests that unknown sequences, similar to genes responsible for the repair of DNA lesions resulting from the alkylation or oxidation of nucleotides, are present in the genome of *M. tuberculosis.*

[009] Thus, there was a need in the art to identify sequences in the genome of *M. tuberculosis* having similarity to genes responsible for the repair of DNA lesions responsible for alkylation or oxidation of nucleotides, such as the *mut* genes from *E. coli.* More particularly, there was a need to identify sequences similar to *E. coli* genes and other *mut* genes. There was a further need to develop methods of using these sequences to predict the epidemic character of a *Mycobacterium tuberculosis* isolate and/or a selective advantage to be maintained in the host and/or the acquisition of multiple drug resistance (MDR) by the isolate.

SUMMARY OF THE INVENTION

[010] Accordingly, this invention aids in fulfilling these needs in the art. More particularly, this invention provides methods for predicting the epidemic character of a *Mycobacterium tuberculosis* isolate and/or a selective advantage to be maintained in the host and/or the acquisition of multiple drug resistance (MDR) by the isolate, wherein the method comprises detecting an alteration in the DNA repair system of the isolate. This invention provides specific pairs of oligonucleotide primers or probes that lead to fragments that hybridize specifically, under stringent hybridization conditions as defined hereinafter, to the nucleic acid (RNA or DNA) from a particular strain of *M. tuberculosis* that has the MDR phenotype. Methods and kits for detecting a strain of *M. tuberculosis* that has the MDR phenotype are also disclosed.

[011] In a first aspect this invention involved an investigation of eleven multidrug resistant *M. tuberculosis* strains isolated in Spain in 1998, 1999, and 2000 (including five Beijing strains), one strain resistant to isoniazid and one strain sensitive to all antibiotics tested. Sensitivity to anti-tuberculosis drugs was assessed by the proportion method. Resistance to rifampicin was confirmed by demonstrating the presence of mutations in *rpoB*. In several cases, streptomycin resistance was confirmed by showing the presence of mutations in the *rpsL* or *rrs* gene. Mutations in putative *mut* genes were looked for by sequencing. The corresponding genetic regions were amplified and directly sequenced with the same primers. The study included data obtained form the published genome sequences of two isolates, H37Rv and CDC1551, which do not have a "Beijing" genotype.

[012] The *mutT2* sequence, the Rv3908 open reading frame (with a *mutT* domain), and *rpoB, rpsL,* and *rrs* sequences were first analyzed in eleven MDR strains (five strains with a "Beijing" genotype and six strains with a genotype other than "Beijing") and in three strains sensitive to all antibiotics tested (one strain with a Beijing genotype and two strains with a genotype other than Beijing). All MDR strains carried mutations in *rpoB.* All MDR strains with a Beijing genotype carried mutations in *rpsL.* All strains with a Beijing genotype (whether antibiotic sensitive or MDR) carried a mutation in Rv3908, a putative *mutT* gene. Three MDR strains with a Beijing genotype carried an additional mutation in *mutT2.*

[013] It was discovered that *M. tuberculosis* strains of the "Beijing" genotype, which have been responsible for several outbreaks of MDR tuberculosis, carry a mutation in a putative *mut* gene. Three MDR strains with a "Beijing" genotype carry an additional mutation in a second putative *mutT* gene. These mutations may have provided these strains with a better adaptability to hostile environments, such that they constitute a higher risk for the patients to develop MDR tuberculosis, especially when these patients receive insufficient anti-tuberculosis treatments.

[014] *M.* tuberculosis sequences similar to *alkA* and *ogt* sequence of *E. coli* were also investigated.

[015] One strain out of six MDR strains with a genotype other than Beijing carry a mutation ACC→AGC at the 15th annotated codon of *ogt*, leading to Thr→Ser.

[016] Two MDR strains with Beijing genotype and one strain with a Beijing phenotype and sensitive to antibiotics carry a mutation of CGC→CTC at the 37th annotated codon of *ogt*, leading to Arg→Leu.

[017] Two MDR Beijing strains carry a silent change at the 12th codon of *ogt*, GGG→GGA.

[018] Four MDR Beijing strains carry a mutation ATC→GTC at the 12th codon of *alkA*, leading to Ile→Val.

[019] Accordingly, this invention provides a method for predicting the epidemic character of a *Mycobacterium tuberculosis* isolate and/or the acquisition of multiple drug resistance (MDR) by the isolate, wherein the method comprises detecting an alteration in the DNA repair system of the isolate. In one embodiment, the isolate contains a mutation in one or more *mutT* locus, and in particular a mutation at one or more mutT family member selected from the Rv3908 locus, the *mutT*2 locus, and the *ogt* locus. In another embodiment, the isolate consists essentially of a Beijing *Mycobacterium tuberculosis* strain. In further embodiments any of the above isolates can also contain a mutation in at least one locus selected from the group consisting of *rpoB*, *rpsl*, *rrs*, or *rpsl*.

[020] To further investigate the correlation between mutations in *mutT* loci and the epidemic character of *M. tuberculosis* isolates, *mut* genes in 170 *M. tuberculosis* complex isolates from 38 different countries, including strains responsible for tuberculosis outbreaks were analyzed. The strains collected in a previous study performed under auspices of the European Concerted Action project on Molecular Epidemiology of Tuberculosis (4) were included in this collection. *M. tuberculosis* strains were previously grouped in 3 classes according to Sreevatsan *et al.* (14). The Beijing strains belong to class 1. Sixty four strains from class 1 including 34 Beijing strains, 47 strains of class 2, three strains of class 3, and 52 strains of an undetermined class were

investigated. Eight Beijing strains were multi-drug resistant (MDR). Sensitivity to anti-tuberculosis drugs was assessed by concentration methods. Resistance to rifampicin was confirmed by demonstrating the presence of mutations in the *rpoB* gene. Mutations in putative *mut* genes were detected by DNA sequencing. Three *mutT*-homologous genes (*mutT1*, *mutT2*, and Rv3908), and *ogt* were examined in 169 strains. Data obtained from the published genome sequences of three isolates, strain MT210, which has a Beijing genotype, and strains H37Rv and CDC1551, which represent other genotypes, were also included.

[021] The vast majority (31 out of 35 strains) of the Beijing strains analyzed carried a mutation replacing Arg 48 by Gly in the deduced polypeptide encoded by the *mutT* putative gene Rv3908. In addition, 23 out of these 31 strains carried another mutation replacing Gly 58 by Arg near the active site of the putative enzyme encoded by *mutT2*. Five of the eight Beijing strains carrying the mutation in Rv3908 but no mutation in *mutT2*, appeared to contain a mutation replacing Arg 37 by Leu in the deduced protein encoded by the putative *ogt* gene. Seven of the eight MDR Beijing strains analyzed contained the mutations in Rv3908 and *mutT2* described above, whereas the eighth had mutations in Rv3908 and *ogt*. None of the 134 non-Beijing strains, representing a variety of genotypes contained mutations in these three genes. The Beijing strain, which was responsible for a recent epidemic in Gran Canaria (5) and which bore the mutations in Rv3908 and *ogt*, was shown to present an increased rate of mutation to rifampicin resistance.

[022] This invention also provides a method for detecting a *Mycobacterium tuberculosis* strain having a multiple drug resistance (MDR) phenotype, wherein the

method comprises detecting a mutation in the Rv3908 locus of the genome of the *Mycobacterium tuberculosis* strain. In one embodiment, the method comprises detecting a mutation at codon 48 of the Rv3908 locus, and in particular, the method comprises detecting GGG at codon 48.

[023] Further, this invention provides a method for detecting a *Mycobacterium tuberculosis* strain having a multiple drug resistance (MDR) phenotype, wherein the method comprises detecting a mutation in the *mutT2* locus of the genome of the *Mycobacterium tuberculosis* strain. In a further embodiment, the method comprises detecting a mutation at codon 58 of the *mutT2* locus, in particular, detecting CGA at codon 58.

[024] In another embodiment, this invention provides a method of detecting a *Mycobacterium tuberculosis* strain having a multiple drug resistance (MDR) phenotype, wherein the method comprises:

(a) providing a biological sample suspected of containing *Mycobacterium tuberculosis;*

(b) amplifying nucleic acids in the sample using as a primer pair

(i) SEQ ID NO: 1, and

SEQ ID NO: 2; or

(ii) SEQ ID NO: 3, and

SEQ ID NO: 4; or

(iii) SEQ ID NO: 5, and

SEQ ID NO: 6; or

(iv) SEQ ID NO: 7, and

SEQ ID NO: 8; and

(c) detecting a mutation in the Rv3908 locus, or the *mutT2* locus, or the *ogt* locus, or the *alkA* locus of the *Mycobacterium tuberculosis*.

[025] In another embodiment, the isolate consists essentially of a Beijing *Mycobacterium tuberculosis* strain. In additional embodiments, the strain contains a mutation in one or more *mutT* family member selected from the Rv3908 locus, the *mutT2* locus, and the *ogt* locus. In another embodiment, the isolate contains a mutation in at least one locus selected from the group consisting of *rpoB*, *rpsl*, *rrs*, or *rpsl*.

[026] In another embodiment, the method comprises detecting a mutation in the Rv3908 locus of the genome of the *Mycobacterium tuberculosis* strain. In a further embodiment, the method comprises detecting a mutation at codon 48 of the Rv3908 locus. And in another embodiment, the method comprises detecting GGG at codon 48.

[027] In another embodiment, the method comprises detecting a mutation in the *mutT2* locus of the genome of the *Mycobacterium tuberculosis* strain. In a further embodiment, the method comprises detecting a mutation at codon 58 of the *mutT2* locus. And in another embodiment, the method comprises detecting CGA at codon 58.

[028] In another embodiment, the method comprises detecting a mutation in the *ogt* locus of the genome of the *Mycobacterium tuberculosis* strain.

[029] This invention also provides a polynucleotide consisting of contiguous nucleotides of the Rv3908 locus of a *Mycobacterium tuberculosis* strain including codon 48 of said locus, or a polynucleotide fully complementary thereto. In one embodiment, codon 48 is GGG. In another embodiment, the polynucleotide contains the complement of SEQ ID NO: 1, SEQ ID NO: 2, or both SEQ ID NOs: 1 and 2.

[030] Further, this invention provides a polynucleotide consisting of contiguous nucleotides of the *mutT2* locus of a *Mycobacterium tuberculosis* strain including codon 58 of said locus, or a polynucleotide fully complementary thereto. In one embodiment, codon 48 is CGA. In another embodiment, the polynucleotide contains the complement of SEQ ID NO: 3, SEQ ID NO: 4, or both SEQ ID NOs: 3 and 4.

[031] In addition, this invention provides a purified polynucleotide comprising a nucleotide sequence selected from:

a) SEQ ID NO: 1;

b) SEQ ID NO: 2;

c) SEQ ID NO: 3;

d) SEQ ID NO: 4;

e) SEQ ID NO: 5;

f) SEQ ID NO: 6;

g) SEQ ID NO: 7; and

h) SEQ ID NO: 8.

[032] Additionally, the invention includes a purified polynucleotide that hybridizes specifically under stringent conditions with one or more polynucleotide sequence selected from SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, and SEQ ID NO: 8.

[033] The invention further includes polynucleotide fragments capable of hybridization under stringent conditions with any one of the nucleotide sequences enumerated above.

[034] Additionally, the invention includes kits for the detection of the presence of *M. tuberculosis* strains that contain the polynucleotide sequences set forth above.

[035] In another embodiment, the invention provides:

a polynucleotide fragment comprising SEQ ID NO: 1 (*mutT2*-1);

a polynucleotide fragment comprising SEQ ID NO: 2 (*mutT2*-2);

a polynucleotide fragment comprising SEQ ID NO: 3 (Rv3908-1);

a polynucleotide fragment comprising SEQ ID NO: 4 (Rv3908-2);

a polynucleotide fragment comprising SEQ ID NO: 5 (*alkA*-1);

a polynucleotide fragment comprising SEQ ID NO: 6 (*alkA*-2);

a polynucleotide fragment comprising SEQ ID NO: 7 (*ogt*-1); and

a polynucleotide fragment comprising SEQ ID NO: 8 (*ogt*-2).

[036] In another embodiment, the invention provides sequences of open reading frames of *M. tuberculosis mut* genes as follows:

a purified polynucleotide of 1488 bp designated as *alkA* and consisting of SEQ ID NO: 27;

a purified polynucleotide of 495 bp designated as *ogt* and consisting of SEQ ID NO: 28;

a purified polynucleotide of 423 bp designated *mutT2* and consisting of SEQ ID NO: 29;

a purified polynucleotide of 744 bp designated Rv3908 and consisting of SEQ ID NO: 30;

a purified polynucleotide of 912 bp designated *mutY* and consisting of SEQ ID NO: 31; and

a purified polynucleotide of 2406 bp designated Rv3909 and consisting of SEQ ID NO: 32.

[037] In another embodiment, the invention provides sequences of cDNA sequences of *M. tuberculosis mut* genes as follows:

a purified polynucleotide comprising SEQ ID NO: 27 (*alkA*);

a purified polynucleotide comprising SEQ ID NO: 28 (*ogt*);

a purified polynucleotide comprising SEQ ID NO: 29 (*mutT2*);

a purified polynucleotide comprising SEQ ID NO: 30 (Rv3908);

a purified polynucleotide comprising SEQ ID NO: 31 (*mutY*); and

a purified polynucleotide comprising SEQ ID NO: 32 (Rv3909).

[038] In another embodiment, the invention provides purified polynucleotide sequences that are delimited upstream by the polynucleotide sequence of an upstream primer and downstream by the polynucleotide sequence of a downstream primer, wherein the purified polynucleotide sequence comprises the open reading frame sequence of a *mut* gene of the invention as follows:

a purified polynucleotide sequence delimited upstream by the polynucleotide sequence of SEQ ID NO: 1 and downstream by the polynucleotide sequence of SEQ ID NO: 2, wherein the purified polynucleotide sequence comprises SEQ ID NO: 29;

a purified polynucleotide sequence delimited upstream by the polynucleotide sequence of SEQ ID NO: 3 and downstream by the polynucleotide sequence of SEQ ID NO: 4, wherein the purified polynucleotide sequence comprises SEQ ID NO: 30;

a purified polynucleotide sequence delimited upstream by the polynucleotide sequence of SEQ ID NO: 5 and downstream by the polynucleotide sequence of SEQ ID NO: 6, wherein the purified polynucleotide sequence comprises SEQ ID NO: 27; and

a purified polynucleotide sequence delimited upstream by the polynucleotide sequence of SEQ ID NO: 7 and downstream by the polynucleotide sequence of SEQ ID NO: 8, wherein the purified polynucleotide sequence comprises SEQ ID NO: 28.

[039] Still further, this invention provides a purified polynucleotide sequence originating from a gene of *M. tuberculosis* comprising a mutator allele.

[040] In another embodiment, the invention provides:

an *E. coli* strain containing the plasmid pMYC2501 deposited at the C.N.C.M. on August 20, 2001 under Accession No. I-2711;

an *E. coli* strain containing the plasmid pMYC2502 deposited at the C.N.C.M. on August 20, 2001 under Accession No. I-2712; and

an *E. coli* strain containing the plasmid pMYC2503 deposited at the C.N.C.M. on August 20, 2001 under Accession No. I-2713.

[041] Still further, this invention provides a method for detecting in a patient infected by *M. tuberculosis* a higher risk of being unable to eliminate the bacillus or a higher probability to develop MDR tuberculosis, wherein the method comprises detecting the presence of mutator alleles in clinical strains of *M. tuberculosis* with one or more polynucleotide fragments selected from SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, and SEQ ID NO: 8.

BRIEF DESCRIPTION OF THE DRAWINGS

[042] This invention will be more fully described with reference to the figures, in which:

[043] Figure 1 depicts an example of the characteristic IS6110 RFLP and highly specific spoligo patterns of the Beijing genotype of *M. tuberculosis*.

[044] Figures 2A-2C: Figures 2A and 2B show the sequence of the *alkA* locus, which is designated SEQ ID NO: 17. The sequences of primers *alkA*-1 (SEQ ID NO: 5) and *alkA*-2 (SEQ ID NO: 6) are indicated by underlining. Figure 2C shows the sequence of the 1488 bp *alkA* open reading frame (ORF) (SEQ ID NO: 27).

[045] Figures 3A-3B: Figure 3A shows the sequence of the *ogt* locus, which is designated SEQ ID NO: 18. The sequences of primers *ogt*-1 (SEQ ID NO: 7) and *ogt*-2 (SEQ ID NO: 8) are indicated by underlining. Figure shows the sequence of the 495 bp *ogt* open reading frame (ORF) (SEQ ID NO: 28).

[046] Figures 4A-4B: Figure 4A shows the sequence of the *mutT2* locus, which is designated SEQ ID NO: 19. The sequences of primers *mutT2*-1 (SEQ ID NO: 1) and *mutT2*-2 (SEQ ID NO: 2) are indicated by underlining. Figure 4B shows the sequence of the 423 bp *mutT2* open reading frame (ORF) (SEQ ID NO: 29).

[047] Figures 5A-5B: Figure 5A shows the sequence of the Rv3908 locus, which is designated SEQ ID NO: 20. The sequences of primers Rv3908-1-1 (SEQ ID NO: 3) and Rv3908-2 (SEQ ID NO: 4) are indicated by underlining. Figure 5B shows the sequence of the 744 bp Rv3908 open reading frame (ORF) (SEQ ID NO: 30).

[048] Figures 6A-6B: Figure 6A shows the sequence of the *mutY* locus, which is designated SEQ ID NO: 21. The sequences of primers *mutY*-1 (SEQ ID NO: 9) and

mutY-2 (SEQ ID NO: 10) are indicated by underlining. Figure 6B shows the sequence of the 912 bp *mutY* open reading frame (ORF) (SEQ ID NO: 31).

[049] Figures 7A-7C: Figures 7A and 7B show the sequence of the Rv3909 locus, which is designated SEQ ID NO: 22. Figure 7C shows the sequence of the 2406 bp Rv3909 open reading frame (ORF) (SEQ ID NO: 32).

DETAILED DESCRIPTION OF THE INVENTION

[050] *M. tuberculosis* strains with a "Beijing" genotype have been associated with outbreaks worldwide, including multidrug resistant (MDR) isolates like the "W" strain in the United States. An aim of this invention was to test the hypothesis of a better adaptation of "Beijing" *M. tuberculosis* due to alterations in DNA repair genes (*mut* genes). Similarly, other MDR strains (with a genotype other than Beijing) may accumulate mutations in DNA or nucleotide repair systems. Indeed, such alterations could result in increased frequencies of mutations in genes responsible for the resistance of the pathogen to different environments as well as to antibiotics, as described for *E. coli* and *Pseudomonas* species. *In silico* analysis of bacterial genomes provides a list of putative *mut* genes. *mut* and *mutT* putative genes in MDR *M. tuberculosis* isolates, including "Beijing" strains, were investigated as follows.

[051] A network to monitor the spread of MDR-Tuberculosis in Spain based on genomic typing was set up in January 1998 by the Spanish Working Group on MDR-TB, coordinated by the Mycobacterial Genetic Unit of the University of Zaragoza and the Instituto de Salud Carlos III. All MDR strains defined as isolates with resistance to at least isoniazid (I) and rifampicin (R) isolated in 1998 in Spain were sent to the University

of Zaragoza for molecular typing by the restriction fragment length polymorphism (RFLP) method using IS6110 (10). A total of 203 strains were registered. Among them 11 were of the Beijing family.

[052] DNA samples of MDR strains were chosen with the criteria of, first exhibiting different restriction fragment length polymorphism (RFLP), and second containing enough DNA for an extensive analysis. In a first set of experiments five "Beijing" MDR strains and six MDR strains with a genotype other than "Beijing" were analyzed. One strain resistant to isoniazid and one strain susceptible to all antibiotics were analyzed, both having a genotype other than "Beijing".

[053] Several putative *mut* genes were annotated as such in the released genome sequence of *M. tuberculosis*. In addition, using a BLAST, Rv3908 was identified as an ORF carrying a *mutT* domain (11, 12).

[054] The following primers were designed to amplify putative *mut* genes:

mutT2:

mutT2-1:

5'-TCCGGATGATGATTTACCTCC-3' SEQ ID NO: 1,

mutT2-2:

5'-TCCGCCGGGTCGGGGAC-3' SEQ ID NO: 2;

Rv3908:

Rv3908-1:

5' -TCGAAGGTGGGCAAATCGTG-3' SEQ ID NO: 3,

Rv3908-2:

5' -TGGGGTTCGCTGGAAGTGG -3' SEQ ID NO: 4;

alkA:

alkA-1*:*

5'-AGCCGCGTAGGTAACCT-3' SEQ ID NO: 5,

alkA-2:

5'-TGCTCGAGCATCCGCAG-3' SEQ ID NO: 6;

ogt:

ogt-1*:*

5'-CAGCGCTCGCTGGCGCC-3' SEQ ID NO: 7,

ogt-2:

5'-GACTCAGCCGCTCGCGA-3' SEQ ID NO: 8;

mutY:

mutY-1:

5' -CCGGCGACGAATCGCTCGTT-3' SEQ ID NO: 9,

mutY-2:

5'-AGCTGGGACAGTCGTCGCGG-3' SEQ ID NO: 10.

mutT1:

5'-ATCGTCGGCGTGCCGTG-3' SEQ ID NO: 23,

5'-GTCAGCGTCCTGCCCGG-3' SEQ ID NO: 24

mutT3:

5'-GTCACGTCTGTTAGGACCTC-3' SEQ ID NO: 25,

5'-CGCGCAACGGCTGCCGG-3' SEQ ID NO: 26

Similarly, primers were designed to amplify:

rpoB:

5'- TACGGTCGGCGAGCTGATCC-3' SEQ ID NO: 11

5'- TACGGCGTTTCGATGAACC-3' SEQ ID NO: 12;

rrs:

5' - GAGAGTTTGATCCTGGCTCAG-3' SEQ ID NO: 13,

5' -TGCACACAGGCCACAAGGGA-3' SEQ ID NO: 14; and

rpsl:

5' -GGCCGACAAACAGAACGT-3' SEQ ID NO: 15,

5' -GTTCACCAACTGGGTGAC-3' SEQ ID NO: 16.

[055] Amplification using primers *mutT2*-1 (SEQ ID NO: 1) and *mutT2*-2 (SEQ ID NO: 2) will result in a polynucleotide sequence delimited upstream by the polynucleotide sequence of SEQ ID NO: 1 and downstream by the polynucleotide sequence of SED ID NO: 2.

[056] Amplification using primers Rv3908-1 (SEQ ID NO: 3) and Rv3908-2 (SEQ ID NO: 4) will result in a polynucleotide sequence delimited upstream by the polynucleotide sequence of SEQ ID NO: 3 and downstream by the polynucleotide sequence of SED ID NO: 4.

[057] Amplification using primers *alkA*-1 (SEQ ID NO: 5) and *alkA*-2 (SEQ ID NO: 6) will result in a polynucleotide sequence delimited upstream by the polynucleotide sequence of SEQ ID NO: 5 and downstream by the polynucleotide sequence of SED ID NO: 6.

[058] Amplification using primers *ogt*-1 (SEQ ID NO: 7) and *ogt*-2 (SEQ ID NO: 8) will result in a polynucleotide sequence delimited upstream by the polynucleotide sequence of SEQ ID NO: 7 and downstream by the polynucleotide sequence of SED ID NO: 8.

[059] DNA sequencing was carried out directly on the amplified fragments by using the dideoxy chain-termination method with the Big Dye Terminator Cycle sequencing Kit (PE Applied Biosystems) on a GeneAmp PCR system 9600; Perkin Elmer) and run on a DNA analysis system model 373 (Applied Biosystems).

[060] Sequences of *mutY*, *mutT2*, Rv3908, *rpoB, mutT1, mutT3, ogt, rrs*, and *rpsl* of strains H37Rv, CDC1551, and MT210 were obtained from published sequences (12) or at the TIGR WEB site. Several putative *mut* genes were annotated as such in the released genome sequence of *M. tuberculosis*. Using a BLAST, Rv3908 was identified as an ORF carrying a *MutT* domain (11, 12). Alleles of putative genes coding for DNA repair enzymes *mutT* (which hydrolyses 8-oxo-deoxyguanosine triphosphate) (13, 14) and *mutY* (specific adenine glycosylase) (8) were investigated in 11 MDR strains of *M. tuberculosis*, and 3 susceptible strains to all antibiotics plus one strain resistant to isoniazid that were taken as controls. In searches for sequences similar to *mut T*, *mutT1*, *mutT2*, *mutT3,* and RV3908 were discovered. The best fits were observed for *mutT2* and Rv3908. The search for sequences similar to *ogt* and *mutY* identified a single open reading frame in each case.

[061] The MDR and antibiotic sensitive strains were identified by classical identification tests. They were isolated from patients in Spain (the European collection of *M. tuberculosis* MDR strains). Five of the ten MDR strains and one sensitive strain

had a "Beijing" genotype. Primers were designed to amplify the *mutY*, *mutT2*, Rv3908, *rpoB, ogt, rrs*, alkA, and *rpsl* gene candidates. For the *mutY* and *mutM* putative genes, amplification was obtained with all strains, but sequencing analysis did not reveal any mutation at these loci.

[062] In all strains with a "Beijing" genotype, whether MDR or antibiotic sensitive, a mutation was discovered in Rv3908. Codon 48 (CGG) of the annotated ORF had been changed to GGG, which would lead to the substitution of Arg by Gly. No mutation in Rv3908 was observed in any other strain with a genotype other than "Beijing" (antibiotic sensitive or MDR).

[063] Three out of four MDR "Beijing" strains had an additional mutation in *mutT2*, which led to a change in codon 58 (GGA to CGA) resulting in a substitution of Gly by Arg. The active site of the *E. coli* MutT enzyme comprises amino acids 53, 56, 57, and 98. Helix I of the polypeptide spans from aa 47 to 59 (15). Therefore, it would seem likely that a Gly→ Arg mutation at position 58 would have a significant effect on enzyme activity and lead to a mutator phenotype, which in turn would facilitate the acquisition of genomic mutations resulting in resistance to antibiotics.
Resistance to rifampicin was confirmed by showing mutations in *rpoB* for all MDR strains. An MDR *M. bovis* strain that was responsible for the B epidemic in Europe was investigated. This strain carries a mutation at the beginning of Rv3909, which seems to be arranged in an operon with RV3908. The different mutations are listed in Table 1.

Table 1: strain characteristics and analysis of mutations at rpoB, rpsl, rrs, Rv3908 and *mut T2*[#]

Beijing MDR strains	Drugs susceptibility R/I/E/S	*rpoB*[a]	*rpsl*[b]	*rrs*[c]	Rv3908	*MutT2*	*alkA*	*ogt*
ZA 20	R/I/S	Asp 516 Val	Lys 43 Arg	wt	Arg 48 Gly	wt	ND	Arg 37 Leu
ZA 65	R/I/E/S	Asp 516 Val	Lys 43 Arg	wt	Arg 48 Gly	wt	Ile 12 Val	Arg 37 Leu
ZA 67	R/I/E/S	His 526 Leu	Lys 43 Arg	wt	Arg 48 Gly	Gly 58 Arg	Ile 12 Val	ND
ZA 68	R/I(at least)	Ser 531 Leu Val 577 Ala	Lys 43 Arg	wt	Arg 48 Gly	Gly 58 Arg	Ile 12 Val	ND
ZA 69	R/I/S	Ser 531 Leu	Lys 43 Arg	wt	Arg 48 Gly	Gly 58 Arg	Ile 12 Val	ND
MDR strains other than Beijing								
ZA 11	R/I/E	Ser 531 Leu	wt	wt	wt	wt	wt	Thr 15 Ser
ZA 12	R/I/E	ND	wt	wt	wt	wt	wt	wt
ZA 13	R/I/S	His 526 Leu	Lys 43 Arg	wt	wt	wt	wt	wt
ZA 14	R/I/E(at least)	His 526 Leu	wt	c491t	wt	wt	wt	wt
ZA 16	R/I/E/S	Ser 531 Leu	wt	wt	wt	wt	wt	wt
ZA 17	R/I/E/S	Asp 516 Val	wt	wt	wt	wt	ND	ND
Beijing strains								
ZA 62	sensitive	ND	ND	ND	Arg 48 Gly	wt	ND	Arg 37 Leu
Strains other than Beijing								
M. bovis MDR ZA19	R/I/E/S/ (at least)	Ser 531 Leu	Lys43Arg	wt	Rv3909	wt	Ile 12 Val	wtt
ZA 15	I/E	wt	wt	wt	wt	wt	wt	wt
CDC1551*	sensitive	wt	wt	wt	wt	wt	wt	wt
H37 Rv**	sensitive	wt	wt	wt	wt	wt	wt	wt

* Analysis of the different loci using the sequences provided by TIGR:

http://www.tigr.org/

** Analysis of the different loci using the sequences provided by SANGER:

http://www.sanger.ac.uk/

R = Rifampin I = Isoniazid E = Ethamutol S = Streptomycin

[a] Gene encoding the RNA polymerase subunit B

[b] Gene encoding ribosomal protein S12

[c] Gene encoding 16S rRNA

[064] # For rpoB, rpsl, Rv3908, alkA, ogt, and mutT2: amino acid substitution; For rrs: nucleotide substitution. It is interesting to note in Table 1 that the three MDR "Beijing" strains with the same mutation at the putative *mut T2* loci harbored a different *rpoB* mutation. The three strains were isolated from patients who had immigrated from Eastern Europe. These findings suggest that the two strains may correspond to the same outbreak. The acquisition of the three different mutations in *rpoB* leading to *rifampicin* resistance (5) must have occurred after the acquisition of mutations in the putative nucleotide repair enzyme genes Rv3908 and *mutT2*.

[065] Resistance to streptomycin correlated with mutations at the *rpsl* locus for all MDR strains with a "Beijing" genotype. Strains with a genotype other than "Beijing" carried mutations at *rrs* (one strain), *rpsl* (one strain), or somewhere else and genetically unidentified (four strains) (4).

[066] In a second set of experiments, *mut* genes in DNA of a total of 170 *M. tuberculosis* complex strains originating from 38 different countries were analyzed. This set comprised 149 *M. tuberculosis* strains, but also included other members of the complex, such as *M. bovis* (12), *M. bovis BCG* (3), *M. africanum* (2), *M. microti* (2), and *M. canettii* (1). Sixty-eight *M. tuberculosis* strains, including eight Beijing genotype strains, of which one was MDR, and the 20 strains of the other species mentioned above were selected, because they were characterized with 13 different genetic markers in previous studies (4, 21).

[067] Furthermore, strains representing different branches of the Beijing genotype were included. Five *M. tuberculosis* strains of the Beijing genotype and one strain of another genotype were obtained from the national program for surveillance of

MDR tuberculosis in Spain. Five *M. tuberculosis* Beijing genotype strains isolated in The Netherlands were included because they exhibited spoligo patterns with fewer than nine spacers. Five additional Beijing genotype strains were included. These strains exhibited hybridization to an additional spacer, as demonstrated using the extended set of spacers (vanSoolingen unpublished), two of which lacked hybridization to spacer. Four Beijing strains representing lineage's of the W-strain and originating in the U.S.A. were also included (22). Five Beijing isolates from patients living in Vietnam were selected because these strains acquired resistance in comparison with initial isolates of those patients. From ten patients, two follow-up isolates were included, of which the second showed an increase in drug resistance, four of these patients originated in Vietnam and six in The Netherlands, one was of the Beijing genotype. Fourteen additional non-Beijing strains were selected from The Netherlands, of which nine contained few IS6110 copies.

[068] Finally, nineteen strains of another frequently observed genotype, the 'Haarlem genotype' (4) were investigated. Five of these originated from Bolivia, two were from an outbreak of MDR-TB in the Czech Republic (23), and twelve were of the "M-type" from Argentina (24).

[069] To summarize, the collection consisted of 34 Beijing genotype isolates, 32 Haarlem genotype isolates, eight strains of the African genotype, four of the Hanoi type and 88 of other genotypes. For the majority of these strains the polymorphism in *katG* and *gyrA*, allowing the grouping according to Sreevatsan et al. 2, were known; 30 strains were of class 1, 47 of class 2, three of class 3 and 86 strains were of an undetermined class.

[070] All isolates were subjected to at least IS6110 RFLP typing and Spoligotyping (4). Drug susceptibility testing according to the proportion method was done for 85 (51%) out of the 166 strains. Eighteen strains were resistant to solely INH, solely streptomycin, or to INH and streptomycin. Twenty strains from Argentina (n=9, all Haarlem), Spain (n=4, all Beijing), Vietnam (n=5, three Beijing, two other), China (n=1, Beijing) and The Netherlands (n=1, other) were MDR. All isolates were unique when the three typing methods were combined.

[071] The sequences of the different genes mentioned above were determined in ten MDR *M. tuberculosis* strains including four Beijing strains. For the *mutY*, *mutT1*, and *mutT3* putative genes, amplification was obtained with all strains tested, but sequence analysis did not reveal any mutation at these loci. However, in comparison with H37Rv and CDC1551, mutations in Rv3908, *mutT2* and *ogt* were observed in MDR Beijing strains, but not in MDR strains with another genotype.

[072] This investigation was extended to look for the presence of mutations in these three genes in the whole collection of *M. tuberculosis* complex isolates. The results are depicted in Tables 2 and 3.

Table 2: Distribution of mutations in the putative genes Rv3908, *mutT2*, and *ogt* among Beijing and non-Beijing genotype *M. tuberculosis* complex strains originating from 38 different countries.

Genotype	Number of strains per affected locus				
	Rv3908	Rv3908 and *mutT2*	Rv3908 and *ogt*	*ogt*	none
Beijing	3[a]	23[b]	5[c]	1[d]	3[e]
Non- Beijing	0	0	0	0	134[f]

[A] *M. tuberculosis* strains from The Netherlands (2) and the USA (1).

[B] *M. tuberculosis* strains from China (1), Malaysia (2), Mongolia (1), The Netherlands (6), South Africa (2), Spain (3) Thailand (1), Vietnam (6); and the genome sequence of *M. tuberculosis* strain MT210 from TIGR.

[C] *M. tuberculosis* strains from The Netherlands (1), South Korea (1), Spain (2), and Vietnam (1).

[d] One *M. tuberculosis* strain from the USA.

[e] *M. tuberculosis* strains from The Netherlands (1) and the USA (2).

[f] 114 *M. tuberculosis* strains from Argentina (14), Bolivia (7), Burundi (2), Canada (2), Central African Republic (2), Chile (2), China (1), Comoro Islands (1), Curacao (1), Czech Republic (4), Ethiopia (1), Ecuador (2), Greenland (2), Honduras (2), India (4), Italy (1), Iran (2), Mongolia (1), The Netherlands (33, including H37Ra), Russia (1), Rwanda (2), South Korea (1), Spain (3), Sri Lanka (2), Tahiti (2), Tanzania (2), Tunisia

(2), Uganda (2), the USA (2), Vietnam (7), Zimbabwe (2), and CDC1551 from TIGR and H37Rv from SANGER ; 12 *M. bovis* strains from Argentina (5), The Netherlands (6), and Saudi Arabia (1); three *M. bovis* BCG strains from Japan (1), The Netherlands (1), and Russia (1); two *M. africanum* strains from the Netherlands; two *M. microti* strains from the UK; and one *M. canettii* strain from Somalia.

Table 3: Affected loci in *M. tuberculosis* strains that are susceptible (S), resistant (DR) or multiresistant (MDR) to antibiotics.

		Affected locus				
Genotype	Drug susceptibility	Rv3908	Rv3908 and *mutT2*	Rv3908 and *Ogt*	*Ogt*	None
Beijing	S	2	12	2	0	1
(n=35)	DR	0	3	1	0	0
	MDR	0	7	1	0	0
	nd	1	1	1	1	2
Non-Beijing	S	0	0	0	0	30
(n=134)	DR	0	0	0	0	14
	MDR	0	0	0	0	12
	nd	0	0	0	0	78

[073] Thirty-one out of 35 strains with a Beijing genotype, either susceptible to tuberculostatics or MDR, had a mutation in Rv3908. Codon 48 (CGG) of the annotated ORF had been changed to GGG, which would lead to the substitution of Arg by Gly (Table 2).

[074] Five out of the 31 Beijing strains with the mutation in Rv3908 carried an additional mutation in *ogt*, implying a change in codon 37, resulting in a substitution of Arg by Leu. In addition, a single Beijing strain carried a mutation in *ogt*, but not in Rv3908 (Table 2).

[075] Twenty-three out of 26 Beijing strains with the mutation in Rv3908, but lacking a mutation in *ogt*, had an additional mutation in *mutT2*, which leads to a change in codon 58 (GGA to CGA) resulting in a substitution of Gly by Arg.

[076] It is noteworthy that all twelve of the MDR Beijing strains carried mutations in two mutT genes. Ten of the twelve carried mutations in Rv3908 and *mutT2*, and the remaining two had a mutation in both Rv3908 and *ogt* (see Table 3).

[077] No mutations in Rv3908, *mutT2*, or *ogt* were observed in any of the 134 *M. tuberculosis* complex strains, originating from 38 different countries, with a genotype other than Beijing, including 30 strains of class 1. Thus, no mutations were observed in any of the 32 strains of the Haarlem genotype, including 14 DR and 12 MDR isolates, nor in any of the strains of *M. bovis*, *M. bovis BCG*, *M. africanum*, *M. microti*, and *M. canettii.*

[078] The sequences of the three DNA repair genes (Rv3908, *mutT2,* and *otg*) were analyzed in 20 serial isolates of ten patients, in which the follow-up isolate showed an increase in resistance to anti-tuberculosis drugs. This analysis revealed that the

sequences of the respective genes were unaltered, irrespective of the genotype. One of these patients was infected with a Beijing genotype strain, and the isolates of this patient had mutations in Rv3908 and *mutT2*. Four of the five additional follow-up isolates of the Beijing genotype that had gained resistance in comparison with previous isolates of the respective patients who suffered from a relapse of tuberculosis also showed mutations in Rv3908 and *mutT*, the fifth showed mutations in Rv3908 and *ogt*.

[079] This invention provides the first demonstration of polymorphism in *M. tuberculosis* genes that might lead to a mutator phenotype, and therefore to a better adaptation of the bacilli to hostile environment (new ref. 28). The vast majority of Beijing strains carried the same mutation in ORF Rv3908, which contains a MutT domain. Seventy-four percent of the Beijing strains carried an additional and identical mutation in a second putative gene of the mutT family, whereas an additional 16% carried an additional and identical mutation in *ogt*. All eight MDR strains with a Beijing genotype were among strains carrying two mutations in putative mutator genes. This would support the notion that M. tuberculosis strains of the Beijing genotype may have adapted to hostile environments, including exposure to anti-tuberculosis drugs, due to a succession of alterations of DNA repair enzymes. Other genes involved in other DNA repair mechanism or in the fidelity of DNA replication may also be involved.

[080] In less than 13% of the Beijing strains no mutation was observed in Rv3908. It is not clear whether these strains are ancestral to the mutator strains, or represent a separate lineage of the Beijing genotype strains. Two of the four Beijing strains devoid of a mutation in Rv3908 were in a separate branch of the dendrogram, when computer-assisted comparison of the IS6110 RFLP patterns using the unweighted

pair group method for clustering was performed. Alternatively, the lack of a mutation in Rv3908 may be due to a reversion that could have occurred after a transient mutator phenotype. The presence of a mutation in *mutT2* was always associated with a mutation in this ORF. This suggests that a first mutation occurred in Rv3908 and that thereafter a second mutation either in *mutT2* or *ogt* was acquired. As observed for other bacterial populations, mutator phenotypes may be transient in many cases to limit deleterious effects (26).

[081] It should be noted that, irrespective of the role that might be played by the mutations described herein, their presence provides a useful marker for predicting the epidemic character of a *Mycobacterium tuberculosis* isolate and/or a selective advantage to be maintained in the host and/or the acquisition of multiple drug resistance (MDR) by the isolate, wherein the method comprises detecting an alteration in the DNA repair system of said isolate.

[082] This invention thus provides specific pairs of oligonucleotide primers or probes that lead to fragments that hybridize specifically, under stringent hybridization conditions as defined hereinafter, to the nucleic acid (RNA or DNA) from a particular strain of *M. tuberculosis* that has the MDR phenotype. These oligonucleotide primers include the following:

(A) SEQ ID NO: 1;

(B) SEQ ID NO: 2;

(C) SEQ ID NO: 3;

(D) SEQ ID NO: 4;

(E) SEQ ID NO: 5;

(F) SEQ ID NO: 6;

(G) SEQ ID NO: 7; and

(H) SEQ ID NO: 8.

[083] In a specific embodiment of the present invention, the purified polynucleotides useful for detecting *M. tuberculosis* strains can be used in combination in order to detect MDR *M.* tuberculosis strains in a biological sample. Thus, the present invention also provides detection methods and kits comprising combinations of the purified polynucleotides according to the invention.

[084] By "polynucleotides" according to the invention is meant the sequences referred to as SEQ ID NOs: 1, 2, 3, 4, 5, 6, 7, or 8 and the complementary sequences and/or the sequences of polynucleotides that hybridize to the referred sequences in specific stringent conditions and that are used for detecting *M. tuberculosis* strains carrying a gene that aids in imparting resistance to antibiotics, such as rifampicin and isoniazid or phenotypes that lead to resistances. The polynucleotides of SEQ ID NOs: 1-8 and their fragments can be used to select nucleotide primers notably for an amplification reaction, such as the amplification reactions further described.

[085] PCR is described in U.S. Patent No. 4,683,202 granted to Cetus Corp. The amplified fragments may be identified by agarose or polyacrylamide gel electrophoresis, or by a capillary electrophoresis, or alternatively by a chromatography technique (gel filtration, hydrophobic chromatography, or ion exchange chromatography). The specificity of the amplification can be ensured by a molecular hybridization using as nucleic probes the entire sequences or the polynucleotides of

SEQ ID NOs: 1-8 and their fragments, oligonucleotides that are complementary to these polynucleotides or fragments thereof, or their amplification products themselves.

[086] Amplified nucleotide fragments are useful as probes in hybridization reactions in order to detect the presence of one polynucleotide according to the present invention or in order to detect the presence of *M. tuberculosis* strains carrying genes that impart resistance to antibiotics. This invention also provides the amplified nucleic acid fragments ("amplicons") defined herein above. These probes and amplicons can be radioactively or non-radioactively labeled, using for example enzymes or fluorescent compounds.

[087] Other techniques related to nucleic acid amplification can also be used and are generally preferred to the PCR technique. The Strand Displacement Amplification (SDA) technique is an isothermal amplification technique based on the ability of a restriction enzyme to cleave one of the strands at a recognition site (which is under a hemiphosphorothioate form), and on the property of a DNA polymerase to initiate the synthesis of a new strand from the 3N OH end generated by the restriction enzyme, and on the property of this DNA polymerase to displace the previously synthesized strand being localized downstream.

[088] The SDA amplification technique is more easily performed than PCR (a single thermostated water bath device is necessary), and is faster than the other amplification methods. Thus, the present invention also comprises using the entire nucleic acid sequences or fragments thereof according to the invention (primers) in a method of DNA or RNA amplification according to the SDA technique.

[089] The entire nucleotide sequences or polynucleotides of SEQ ID NOs: 1-8 and their fragments, especially the primers according to the invention, are useful as technical means for performing different target nucleic acid amplification methods such as:

- TAS (Transcription-based Amplification System), described by Kwoh et al. (27);

- SR (Self-Sustained Sequence Replication), described by Guatelli et al. (28);

- NASBA (Nucleic acid Sequence Based Amplification), described by Kievitis et al., 1991 (29); and

- TMA (Transcription Mediated Amplification).

[090] The polynucleotides of SEQ ID NOs: 1-8 and their fragments, especially the primers according to the invention, are also useful as technical means for performing methods for amplification or modification of a nucleic acid used as a probe, such as:

- LCR (Ligase Chain Reaction), described by Landegren et al., 1988 (30), and improved by Barany et al. (31), who employ a thermostable ligase;

- RCR (Repair Chain Reaction), described by Segev, D, et al. (32);

- CPR (Cycling Probe Reaction), described by Duck et al. (33); and

- Q-beta replicase reaction, described by Miele et al. in 1983 and improved by Chu et al. in 1986, Lizardi et al. in 1988, and by Burg et al. and Stone et al. in 1996.

[091] When the target polynucleotide to be detected is RNA, for example mRNA, a reverse transcriptase enzyme can be used before the amplification reaction in order to obtain a cDNA from the RNA contained in the biological sample. The generated cDNA can be subsequently used as the nucleic acid target for the primers or

the probes used in an amplification process or a detection process according to the present invention.

[092] Nucleic acid probes according to the present invention are specific to produce an amplicon of the invention. By "specific probes" according to the invention is meant any oligonucleotide that hybridizes with one polynucleotide of SEQ ID NOs: 1-8 and which does not hybridize with unrelated sequences.

[093] In a specific embodiment, the purified polynucleotides according to the present invention encompass polynucleotides having at least 75% homology in their entire nucleotide sequences or nucleic acid sequences with SEQ ID NOS: 1-8. By percentage of nucleotide homology according to the present invention is intended a percentage of identity between the corresponding bases of two homologous polynucleotides, this percentage of identity being purely statistical and the differences between two homologous polynucleotides being located at random and on the whole length of said polynucleotides.

[094] The percent identity can be determined, for example, by comparing sequence information using the GAP computer program, version 6.0 described by Devereux et al. (34) and available from the University of Wisconsin Genetics Computer Group (UWGCG). The GAP program utilizes the alignment method of Needleman and Wunsch (35), as revised by Smith and Waterman (36). The preferred default parameters for the GAP program include: (1) a unary comparison matrix (containing a value of 1 for identities and 0 for non-identities) for nucleotides, and the weighted comparison matrix of Gribskov and Burgess (37) as described by Schwartz and

Dayhoff, (38); (2) a penalty of 3.0 for each gap and an additional 0.10 penalty for each symbol in each gap; and (3) no penalty for end gaps.

[095] The oligonucleotide probes according to the present invention hybridize under stringent conditions with *M.* tuberculosis DNA and RNA. As an illustrative embodiment, the stringent hybridization conditions used in order to specifically detect *M. tuberculosis* strains according to the present invention are advantageously the following:

[096] Prehybridization and hybridization are performed at 68EC in a mixture containing:

- 5X SSPE (1X SPE is 0.18 M NaCl, 10mM NaH_2PO_4);
- 5X Denhardt's solution;
- 0.5% (w/v) sodium dodecyl sulfate (SDS); and
- 100 µg ml^{-1} salmon sperm DNA.

[097] The washings are performed as follows:

(a) Two washings at laboratory temperature for 10 min. in the presence of 2X SSPE and 0.1% SDS;

(b) One washing at 68EC for 15 min. in the presence of 1X SSPE and 0.1% SDS; and

(c) One washing at 68EC for 15 min. in the presence of 0.1X SSPE and 0.1% SDS.

[098] The non-labeled polynucleotides or oligonucleotides of the invention can be directly used as probes. Nevertheless, the polynucleotides or oligonucleotides can be generally labeled with a radioactive element (^{32}P, ^{35}S, ^{3}H, ^{125}I) or by a non-isotopic molecule (for example, biotin, acetylaminofluorene, digoxigenin, 5-bromodeoxyuridin,

fluorescein) in order to generate probes that are useful for numerous applications. Examples of non-radioactive labeling of nucleic acid fragments are described in the French Patent No. FR 78 10975 or by Urdea et al. (39) or Sanchez-Pescador et al. (40).

[099] Other labeling techniques can also be used, such as those described in the French patents 2 422 956 and 2 518 755. The hybridization step may be performed in different ways (Matthews et al. 1988). A general method comprises immobilizing the nucleic acid that has been extracted from the biological sample on a substrate (nitrocellulose, nylon, polystyrene) and then incubating, in defined conditions, the target nucleic acid with the probe. Subsequent to the hybridization step, the excess amount of the specific probe is discarded, and the hybrid molecules formed are detected by an appropriate method (radioactivity, fluorescence, or enzyme activity measurement).

[0100] In another advantageous embodiment of the present invention, the probes described herein can be used as "capture probes", and are for this purpose immobilized on a substrate in order to capture the target nucleic acid contained in a biological sample. The captured target nucleic acid is subsequently detected with a second probe, which recognizes a sequence of the target nucleic acid that is different from the sequence recognized by the capture probe.

[0101] The oligonucleotide fragments useful as probes or primers according to the present invention can be prepared by cleavage of the polynucleotides with restriction enzymes, as described in Sambrook et al. in 1989. Another appropriate preparation process of the nucleic acids of the invention containing at most 200 nucleotides (or 200 bp if these molecules are double-stranded) comprises the following steps:

- Synthesizing DNA using the automated method of beta-cyanethylphosphoramidite described in 1986;

- cloning the thus obtained nucleic acids in an appropriate vector; and

- purifying the nucleic acid by hybridizing to an appropriate probe according to the present invention.

A chemical method for producing the nucleic acids according to the invention, which have a length of more than 200 nucleotides (or 200 bp if these molecules are double-stranded), comprises the following steps:

- Assembling the chemically synthesized oligonucleotides having different restriction sites at each end;

- cloning the thus obtained nucleic acids in an appropriate vector; and

- purifying the nucleic acid by hybridizing to an appropriate probe according to the present invention.

[0102] The oligonucleotide probes according to the present invention can also be used in a detection device comprising a matrix library of probes immobilized on a substrate, the sequence of each probe of a given length being localized in a shift of one or several bases, one from the other, each probe of the matrix library thus being complementary to a distinct sequence of the target nucleic acid. Optionally, the substrate of the matrix can be a material able to act as an electron donor, the detection of the matrix positions in which hybridization has occurred being subsequently

determined by an electronic device. Such matrix libraries of probes and methods of specific detection of a target nucleic acid are described in the European patent application No. 0 713 016, or PCT Application No. WO 95 33846, or also PCT Application No. WO 95 11995 (Affymax Technologies), PCT Application No. WO 97 02357 (Affymetrix Inc.), and also in U.S. Patent No. 5,202,231 (Drmanac), said patents and patent applications being herein incorporated by reference.

[0103] The polynucleotide probes according to the invention can be incorporated into a kit for detecting the presence of *Mycobacterium tuberculosis*, wherein the kit comprises:

(A) a polynucleotide probe according to the invention; and

(B) reagents necessary to perform a nucleic acid hybridization reaction.

[0104] In one embodiment, the *Mycobacterium tuberculosis* is present in a biological sample, has resistance to antibiotics, and harbors one or more of the polynucleotide sequences according to the invention.

[0105] The primer pairs according to the invention can be incorporated into another kit for detecting the presence of *Mycobacterium tuberculosis*, wherein the kit comprises:

(A) a primer pair according to the invention; and

(B) reagents necessary to perform a nucleic acid amplification reaction.

[0106] In one embodiment, the *Mycobacterium tuberculosis* is present in a biological sample, has resistance to antibiotics, and harbors one or more of the polynucleotide sequences according to the invention.

[0107] In summary the results described above show that MDR strains of *M. tuberculosis* acquired mutations expected to result in a mutator phenotype. This invention provides the first evidence of polymorphism in putative genes of *M. tuberculosis* that correlate with a mutator phenotype, and therefore to a better adaptation of the bacilli to hostile environment (16). Results presented here show that all Beijing strains had a mutation in the ORF of Rv3908, which contains a MutT domain. Three of these Beijing strains, which also had an MDR phenotype, harbored an additional and identical mutation in a second putative gene of the *mutT* family. This suggests that *M. tuberculosis* strains adapt to hostile environments by a succession of alterations of DNA repair enzymes.

[0108] The acquisition of mutator alleles was described as an adaptive response of bacteria to a succession of different environments (17, 18). After infecting a host, *M. tuberculosis* has to adapt to different environments such as alveolar macrophages, then to granuloma containing unactivated macrophages, and then to activated macrophages after induction of the acquired immune responses. In addition, the bacilli have to adapt to the caseous media with low oxygen concentration in the center of tubercles, and then to different tissues during dissemination of the disease. Such variable growth conditions might select for mutator mutations in *M. tuberculosis* strains as described in other bacterial populations submitted to different environments. Mutations might occur with an increased frequency due to the toxic radicals produced in phagocytic cells. The

accumulation of mutations leading to antibiotic resistance may be a consequence of the appearance of a mutator phenotype. MDR strains would be easily selected when patients with better adapted strains received inadequate anti-tuberculosis regimen.

[0109] An identification of the presence of mutator alleles in clinical strains allows one to identify the patients who present a higher risk of being unable to eliminate the bacillus or to develop MDR tuberculosis, and encourage clinicians to increased vigilance.

[0110] Plasmids containing polynucleotides of the invention have been deposited at the Collection Nationale de Cultures de Microorganismes ("C.N.C.M.") Institut Pasteur, 28, rue du Docteur Roux, 75724 Paris Cedex 15, France, as follows:

Plasmid	Accession No.	Deposit Date
pMYC2501	I-2711	August 20, 2001
pMYC2502	I-2712	August 20, 2001
pMYC2503	I-2713	August 20, 2001

A copy of the deposit receipt for each plasmid is attached hereto, and the entire contents of each deposit receipt are incorporated by reference herein.

REFERENCES

[0111] The following publications have been cited herein. The entire disclosure of each publication is relied upon and incorporated by reference herein.

1 Farmer P, Kim JY. Community based approaches to the control of multidrug resistant tuberculosis: introducing "DOTS-plus". *BMJ* 1998; 317: 671-674.

2 Telzak EE, Sepkowitz K, Alpert P, Mannheimer SH, Medard F, El-Sadr W, Blum S, Gagliardi A, Salomon N, Turett G. Multi-Resistant Tuberculosis in patients without HIV infection. *NEJM* 1995; 333: 907-912.

3 Bifani PJ. Identification of a W variant outbreak of Mycobacterium tuberculosis via population-based molecular epidemiology. *JAMA* 1999; 282(24): 2321-7.

4 Kremer K, van Soolingen D, Frothingham R, et al. Comparison of methods based on different molecular epidemiological markers for typing of Mycobacterium tuberculosis complex strains: interlaboratory study of discriminatory power and reproducibility. *Journal of Clinical Microbiology* 1999; 37(8): 2607-18.

5 Caminero JA. Epidemiological evidence of the spread of a Mycobacterium tuberculosis strain of the Beijing genotype on Gran Canaria Island. *American Journal of Respiratory & Critical Care Medicine* 2001; 164(7): 1165-70.

6 Pfyffer GE. Multidrug-resistant tuberculosis in prison inmates, Azerbaijan. *Emerging Infectious Diseases* 2001; 7(5): 855-61.

7 Van Rie A, Warren R, Mshanga I, Jordaan AM, Gian D, van der Spuy GD, Richardson M, Simpson J, Gie RP, Enarson DA, Beyers N, van Helden PD, Victor TC. Analysis for a limited number of gene codons can predict drug resistance of *Mycobacterium tuberculosis* in a high-incidence community. *J. Clin Microbiol* 2001; 39: 636-641.

8 Finken M, Kirschner P, Meier A, Wrede A, Bottger EC. Molecular basis of streptomycin resistance in *Mycobacterium tuberculosis*: alteraions of the ribosomal protein S12 gene and point mutations within a functional 16S ribosomal RNA pseudoknot. *Mol. Microbiol.* 1993; 9: 1239-1246.

9 Telenti A, Imboden P, Marchesi F, Lowrie D, Cole S, Colston MJ, Matter L, Schopfer K, Bodmer T. Detection of rifampicin- resistance mutations in *Mycobacterium tuberculosis. Lancet* 1993; 341: 647-650.

10 Oliver A, Canton R, Campo P, Baquero F, Blazquez J. High frequency of hypermutable *Pseudomonas aeruginosa* in Cystic Fibrosis Lun Infection. *Science* 2000; 288: 1251-1253.

11 Giraud A, Matic I, Tanillon O, Clara A, Radman M, Fons M, Taddei F. Costs and benefits of high mutation rates: Adaptive evolution of bacteria in the mouse gut. *Science* 2001; 291: 2606-2608.

12 Horst JP, Wu TH, Marinus MG. Echerichia coli mutator genes. *Trends Mirob* 1999; 7: 29-36.

13 Mizrahi V, Andersen SJ. DNA repair in Mycobacterium tuberculosis. What have learnt from the genome sequence? *Mol Microbiol* 1998; 29: 1331-1339.

14 Sreevatsan S. Restricted structural gene polymorphism in the Mycobacterium tuberculosis complex indicates evolutionarily recent global dissemination. *PNAS USA* 1997; 94(18): 9869-74.

15 Samper S, Iglesias MJ, Tello O. The Spanish multidrug resistant tuberculosis network. *Eurosurveillance* 2000; 5: 43-45.

16 Altschul SF, Madden TL, Schaffer, AA, Zhang J, Zhang Z, Miller W, Lipman DJ. Gapped Blast and PSI Blast: a new generation of protein database search programs. *Nucleic Acid Res* 1997; 25: 3389-3402.

17 Cole ST, Brosch R, Parkhill J, Garnier T, Churcher C, Harris D, Gordon SV, Eiglmeier K, Gas S, Barry III CE, Tekaia F, Badcock K, Basham D, Brown D, Chillingworth T, Connor R, Davies R, Devlin K, Feltwell T, Gentles S, Hamlin N, Holroyd S, Homsby T, Jagles K, Kroghs A, Mclean J, Moule S, Morphy L, Oliver K, Osborne J, Quail MA, Rajandream MA, Rogers J, Rutter S, Seeger K, Skelton J, Squares R, Squares S, Sulston JE, Taylor K, Whitehead S, Barrell BG. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* 1998; 393: 537-544.

18 Taddei F, Hayakawa H, Bouton MF, Cirinesi AM, Matic I, Sekiguchi M, Radman M. Counteraction by MutT protein of transcriptional errors caused by oxidative damage. *Science* 1997; 278: 128-130.

19 Cunningham RP. DNA repair: Caretakers of the genome? *Curr Biol* 1997; 7: 576-579.

20 Harris TK. Wu G, Massiah MA, Mildvan AS. Mutational, kinetic, and NMR studies of the roles of conserved glutamate residue and of Lysine-39 in the mechanism of the mutt pyrophosphohydrolase. *Biochemistry* 2000; 39: 1655-1674.

21 Supply P, Lesjean S, Savine E, Kremer K, van Soolingen D, Locht C. Automated high-throughput genotyping for study of global epidemiology of Mycobacterium tuberculosis based on mycobacterial interspersed repetitive units. *Journal of Clinical Microbiology* 2001; 39(10): 3563-71.

22 Kurepina NE, Sreevatsan S, Plikaytis BB, et al. Characterization of the phylogenetic distribution and chromosomal insertion sites of five IS6110 elements in Mycobacterium tuberculosis: non-random integration in the dnaA-dnaN region. *Tubercle & Lung Disease* 1998; 79(1): 31-42.

23 Kubin M, Havelkova M, Hyncicova I, et al. A multidrug-resistant tuberculosis microepidemic caused by genetically closely related Mycobacterium tuberculosis strains. *Journal of Clinical Microbiology* 1999; 37(8): 2715-6.

24 Ritacco V, Di Lonardo M, Reniero A, et al. Nosocomial spread of human immunodeficiency virus-related multidrug-resistant tuberculosis in Buenos Aires. *Journal of Infectious Diseases* 1997;176(3):637-42.

25 Tenaillon O. Mutators, population size, adaptive landscape and the adaptation of asexual populations of bacteria. *Genetics* 1999; 152(2): 485-93.

26 Taddei F. Role of mutator alleles in adaptive evolution. *Nature* 1997; 387(6634): 700-2.

27 Kwoh et al., *PNAS USA* 1989; 86: 1173-77.

28 Guatelli et al., *PNAS USA* 1990; 87: 1874-78.

29 Kievitis et al., *J. Virol. Methods* 1991; 35: 273-86.

30 Landegren et al., *Science* 1988; 241: 1077-80.

31 Barany et al. *PNAS USA*; 88: 189-93.

32 Segev, D, et al., in "Non-radioactive Labeling and Detection of Biomolecules," Keasler, C., Springer Verlag, Berlin, N. Y., pp. 197-205.

33 Duck et al., *Biotechnique* 1990; 9: 142-47.

34 Devereux et al., *Nucl. Acids Res.* 1984; 12: 387.

35 Needleman and Wunsch, *J. Mol. Biol.* 1970; 48:433.

36 Smith and Waterman, *Adv. Appl. Math* 1981; 2: 482.

37 Gribskov and Burgess, *Nucl. Acids Res.* 1986; 14: 6745.

38 Schwartz and Dayhoff, eds., Atlas of Protein Sequence and Structure, National Biomedical Search Foundation, pp. 353-358, 1979.

39 Urdea et al., *Nucleic Acids Research* 1988; 11: 4937-57.

40 Sanchez-Pescador et al., *J. Clin. Microbiol.* 1988; 26(10): 1934-38.

41 Tenaillon O, Toupance B, Nagard HL, Taddei F, Godelle B. Mutators, population size, adaptive landscape and adaptation of asexual populations of bacteria. *Genetics* 1999; 152: 485-493.

42 Rainey PB, Moxon ER. When being hyper keeps you fit. *Science* 2000; 288: 186-1187.

43 Taddei F, Radman M, Smith JM, Toupance B, Gouyon PH, Godelle B. Role of mutator alleles in adaptive evolution. *Nature* 1997; 387: 700-702.